This paper was presented as part of the main technical program at IEEE INFOCOM 2011

Location Privacy Protection from RSS Localization System Using Antenna Pattern Synthesis

Ting Wang
Department of Electrical and Computer Engineering
Virginia Polytechnic Institute and State University
Email: wangting@vt.edu

Yaling Yang
Department of Electrical and Computer Engineering
Virginia Polytechnic Institute and State University
Email: yyang8@vt.edu

***Abstract*—This paper studies the problem of location privacy protection in wireless LAN (WLAN) environment, where received signal strength (RSS) at access points (AP) can potentially be obtained by adversaries to obtain the location of a legitimate mobile station. We propose a two-step location privacy protection scheme using a linear smart antenna array on the mobile station. In the first step, the mobile station observes the arrangement of surrounding APs by moving around and estimating the path losses from itself to the APs. Based on the path loss information, in the second step, the mobile station optimizes the radiation pattern of its smart antenna so that its location privacy is protected while its communication quality is not affected. Two strategies are used in the radiation pattern optimization. The first strategy is to limit the number of APs in range of the mobile station to a safe level so that there are not enough measurements from the APs to make an estimation of the mobile station's location. If the first strategy is not possible, the mobile station falls to the second strategy, where its radiation pattern introduces maximum bias to any location estimation attempt so that the mobile station's true location is not revealed. Simulation results show that compared with traditional transmit power control (TPC) scheme, the first strategy significantly increases the probability of inadequate measurements for location computation. Simulation also demonstrates that the second strategy can significantly degenerate the precision of the positioning system. In many cases, the degenerated location precision is as low as the coverage range of the AP that the mobile station is associated with for communications. This essentially means that the second strategy can invalidate the use of RSS measurement for precise localization.**

I. Introduction

While wireless localization techniques enable the mobile users to enjoy location based services in pervasively deployed WiFi networks, they also threaten the location privacy of these mobile users. While surfing online with a mobile device, a network user can be physically localized and tracked by malicious parties who get control over the localization system or get access to information that can be used to derive the location of the mobile devcie [1], [2].

The level of threat to location privacy depends on the positioning techniques, which can be grouped into three categories according to who is carrying out the location estimation process. In the first category, the mobile station executes self-localization based on the information provided by the network infrastructure [3]. There is less privacy issue in this scenario as long as the mobile station does not report its location to trustless parties. In the second category, the localization systems need the mobile station's cooperation (e.g. reporting transmit power or RSS readings to the positioning system) for location computation [4]. It is possible for the mobile station to choose not to cooperate with the localization system or provide fuzzy information to the localization system whenever it is unwilling to reveal its location. Hence, the location privacy is under the control of the mobile user. In the third category, called passive localization, a mobile station is localized by either the network provider or a third party that has access to the necessary information [5], [6]. This type of localization schemes analyze the mobile station's signal over the air and can localize the mobile station as long as it emits communication signals. Since they do not require cooperation of the mobile station, it is most challenging to defend the location privacy of a mobile user from these localization systems because emitting signal is inevitable during wireless communications. Under such circumstances, the mobile station losses control of its location privacy and is in danger of getting its location information exposed to ill-disposed parties. Hence, location privacy of a mobile user is most threatened by passive localization schemes. Existing methods [7] that intent to protect user location privacy from this type of localization systems forces mobile stations to shutdown its communications for long period of time and hope the mobility of the mobile station during the silent period can make it difficult for the localization system to track the mobile station. However, the long interruptions to communications are highly undesirable and a mobile station may not want to constantly move in practice. Hence, we believe that none of the existing location privacy scheme is able to protect mobile users from the third type of localization schemes.

The aim of this paper, hence, is to address this challenging open problem of protecting location information of a mobile user from passive localization systems. Specifically, we focus on defending user location privacy from received signal strength (RSS) based localization scheme. Although besides RSS, time of arrival (TOA) and angle of arrival (AOA) are also used by current physical layer localization algorithms, RSS-based localization is most often adopted because it requires no extra hardware support and provides acceptable accuracy in WLAN environment. Because of the popularity of RSS-based localization systems, we only focus on protecting location privacy against them in WLAN.

The main contribution of this paper is that it solves the location privacy problem in physical layer through antenna pattern

synthesis. To protect location information in physical layer, we use a linear smart antenna array to change the mobile station's radiation pattern from omnidirectional into an optimized radiation pattern. The radiation pattern reduces the number of valid RSS measurements that can be obtained by the localization system and breaks the trilateration principle among the RSS measurements. Without enough measurements, the localization system cannot uniquely localize the mobile station. Even if enough RSS measurements are gathered, the localization result will contain large bias caused by the irregular radiation pattern of the mobile station. At the same time, our radiation pattern synthesize scheme ensures that the mobile station's communications are intact and we do not require the mobile station to move. To our best knowledge, our work is the first application of pattern synthesis for the purpose of protecting location privacy of wireless mobile users.

The rest of the paper is organized as follows. In Section II, we briefly summarize related works in the area of localization and location privacy. Section III provides the location privacy threat model, a brief overview of the proposed scheme and the antenna model used in this paper. The detailed introduction of the proposed scheme is provided from Section IV to Section VI. The simulation results are discussed in Section VII and the last section concludes our work.

II. Related Work

In this section, we first briefly introduce typical RSS-based localization methods and then provide an overview of current works about location privacy protection.

A. RSS Localization techniques

Current RSS localization techniques can be generally grouped into fingerprint based approaches [5], [8], [9] and propagation model based approaches [3], [6], [10]. Fingerprint localization usually consists of an off-line phase and an on-line phase. Before the positioning system can operate, the off-line phase is required to collect RSS measurements at known locations and a database is built up for pattern matching. During the on-line phase, the actual RSS measurements of the target mobile station is compared with the stored database to return a location estimation. Fingerprint of wireless signal highly depends on the specific environment and is not transportable to different places. Consequently, for every different interested area, the off-line phase must be conducted from the very beginning. Additionally, any change of the infrastructure distribution and the physical environment will necessitate an update of the localization system and the collection of new training data. The disadvantage of fingerprint localization is that it requires lots of human work and is time consuming.

In propagation based RSS localization schemes, large scale path loss is related to the distance between the transmitter and the receiver. Obstacles and noise can also be taken into account in the model. Then given the path loss from the transmitter to the receiver, the distance between them can be estimated and the location of the transmitter can be computed.

B. Location privacy schemes

Four types of location privacy techniques have been proposed to protect location privacy at system level in location-based services (LBS) [11]. These techniques are named "policies", "modification of request", "dummy requests" and "provider change". Policy approaches restrict the precision and conditions under which the location-based service provider (LBSP) can obtain a certain station's location information. Modification of request approaches protect location information by either hiding the user's identifier to the LBSP, or reporting indefinite location information to the LBSP. Dummy requests are designed to confuse the location attacker by generating simulated user requests. Provider change refers to frequently changing the LBSP for a certain service. Recently Meyerowitz and Choudhury [12] developed a new type of camouflage system named "CacheCloak" as a trusted server. It submits intersecting predicted paths of multiple users to the LBSP. Thus the LBSP won't be able to track the path of any individual user. All these existing works discussed so far study how to keep location information unrevealed to LBSP, and some of them depend on a trustable third party server. Thus these methods are not suitable for protecting physical level location privacy in WLAN, where the signal transmitted by the mobile station can reveal its location information.

There are two pioneering works that attempt to protect location privacy from passive localization in physical layer. In [13] Jiang et al. use intelligent transmit power control (TPC) to reduce the APs in range in order to reduce the chance of being localized. While the scheme is simple to implement, its effectiveness is limited when the density of APs in range is high. In [14], a framework of physical layer location privacy scheme with beamforming in AOA localization systems is proposed, where the aim is to reduce the number of possible AOA measurements to thaw any positioning attempt. The authors assume that "signal-to-noise-ratios for successful direction-finding are more stringent than those required for mere communications". They claim that within a distance range the mobile station is able to communicate with the AP, while the AP cannot estimate the AOA of the mobile station's signal. The problem with this work is that its assumption that communication requires lower SNR than direction finding is questionable. However, we acknowledge that this work inspires us with the idea of using intelligent antenna radiation pattern to protect location information in physical layer.

III. Scheme Overview

A. Threat model

In this paper, we focus on protecting location privacy from potential RSS-based location attacks. Although localization algorithms based on AOA or TOA may have better accuracy, RSS-based localization is more likely to be used by an adversary because it requires no special hardware and, hence, is easy to implement in normal WLAN.

In our threat model, off-the-shelf APs in a legitimate WLAN act as anchors and provide RSS measurements at different known locations. We assume that an adversary can remotely

tap into the software systems of these APs and obtain the RSS information. The attacker feeds the RSS information to RSS localization algorithm to localize and track a mobile user.

We do not consider the case that the adversary plants his/her own measurement anchors. It is very expensive, risky and, hence, unlikely for the adversary to physically planting a large number of his/her own private APs to track a mobile user. Directly hacking into the existing WLAN infrastructures is a better and more likely choice for the adversary.

B. Overview of our privacy protection scheme

With location privacy concern, the mobile user's goal is to make the localization system unable to correctly localize him/her while maintaining reliable access to the wireless network. Our scheme realizes this goal by equipping the mobile user with a smart antenna and tuning the radiation pattern of the antenna. This scheme is feasible since it has been experimentally proved that attenuation and amplification of the RSS measurements collected by some APs (anchors of the localization system) can significantly degenerate the performance of a RSS-based localization system [15], [16]. The signal attenuation and amplification can be carried out either at the transmitter or at the receiver. These facts make radiation pattern synthesis a good choice to protect location privacy in physical layer. By changing the antenna radiation pattern, a mobile station can reduce the number of APs that can monitor its signal and degenerate the precision of the adversary's localization system while keeping good communication quality with its associated AP. Meanwhile, since radiation pattern synthesis is performed at the transmitter end, it does not hurt the performance of the WLAN to other mobile users.

In our design, the first strategy for the mobile station is to tune the antenna pattern so that the number of APs that can detect the mobile station's signal is minimized. The objective of this strategy is to ensure that not enough RSS measurements can be obtained for a unique location estimation. However, completely preventing the number of RSS measurements to go beyond a safe level is not always possible when there is uncertainty in signal attenuation and the density of APs is high. When a mobile station finds out that it is unlikely to limit the number of APs that can hear itself to a safe level, the mobile station uses the second strategy, which tries to maximize the bias in the location estimation result. The larger the bias is, the less probable that the true location of the mobile station is learned by the adversary. These two strategies constitute our location privacy scheme.

Fig. 1 illustrates an overview of the whole scheme. The first step of our scheme is silently observing the beacon signals of the surrounding APs and estimating the locations of the APs and the path losses from the mobile station to the APs. Next, based on the estimated AP locations, the mobile station checks if it can use the first strategy to limit the number of APs that can hear its signal to be less than 4. This is because for 2-D localization based on RSS, 4 RSS measurements are needed when the transmit power of the mobile station is unknown to the location system. If this is not possible, the mobile station switch to the second strategy. By optimizing the radiation pattern based on the estimated AP locations, the mobile user introduces various attenuation and amplification in the RSS measurements to degenerate the localization performance.

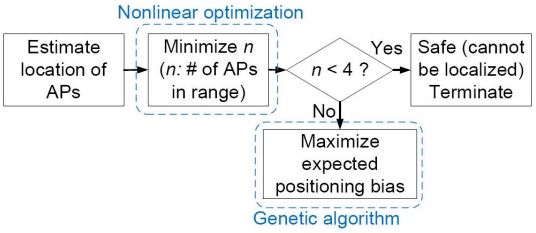


Fig. 1. Scheme overview.

C. Smart antenna model

In our location privacy strategies introduced in the last subsection, either limiting the number of RSS measurements or distorting the RSS measurements requires that the mobile station can control its radiation pattern and transmit power. Hence, how to control the mobile station's radiation pattern and transmit power is the first key problem that need to be addressed. In this paper, we make use of radiation pattern synthesis over smart antenna to achieve our location privacy scheme. Pattern systhesis enables the mobile station to control the signal strength in different directions by tuning the complex weight vector in the beamforming function of an antenna.

In the remainder of this paper, we assume that the mobile station is equipped with a linear smart antenna array with N_{ant} isotropic antenna elements uniformly spaced at distance l. This smart antenna model is also called uniform N_{ant}-element linear antenna array. The beamforming function of this antenna is given by [17] as:

$$G(\theta) = \sum_{i=1}^{N_{ant}} w_i \exp(-j\frac{2\pi}{\lambda} l \cos\theta), \tag{1}$$

where λ is the signal wavelength, θ represents the direction and $\mathbf{w} = [w_1, w_2, \cdots w_{N_{ant}}]^T$ is the complex weight vector which can be designed to change the radiation pattern.

It is important to note that we choose this N_{ant}-element linear antenna array as an example to illustrate the design of our scheme. Our scheme can also work with antennas with other geometric forms, such like circular arrays, planar arrays, and conformal arrays. Although their beamforming functions differ from equation (1), since their radiation patterns are also determined by the complex weight vectors, they can still work with our scheme by simply replacing equation (1) with their corresponding beamforming functions.

IV. Passive Estimation of the Locations of Surrounding APs

In order to determine the optimal radiation pattern, a mobile station needs to know its neighboring APs' locations and the path losses between itself and the APs. However such information is hardly exposed to an ordinary mobile station. In our scheme, we leverage the fact that in WLANs, APs periodically send out beacon signals to announce their existence

to mobile stations. Hence, by intelligently measuring APs' beacon signal, we can design a passive estimation method to get APs' location information. This passive scheme has three steps. In the first step, the mobile station passively measures RSS of the beacon signals from APs at several locations. In the second step, the mobile station estimates APs' locations based on the gathered information. In the third step, path losses to the APs are computed. Following is the detailed description of these three steps.

A. Passive RSS measurement of AP beacon signals

Assume that a mobile station has a desired location, called its desired communication spot, where it wants to conduct its communications in the WLAN. To protect its own privacy, before the mobile station starts its wireless communications, it first starts a process called listen-only wardriving [18] as shown by Fig. 2. In this wardriving process, the mobile station selects 4 different locations, including its desired communication spot, in its neighboring area. Equipped with passive wardriving software, e.g. Kismet [19], and a GPS receiver, the mobile station measures the beacon signal strength of the surrounding APs at the selected locations and record the coordinates of these locations.

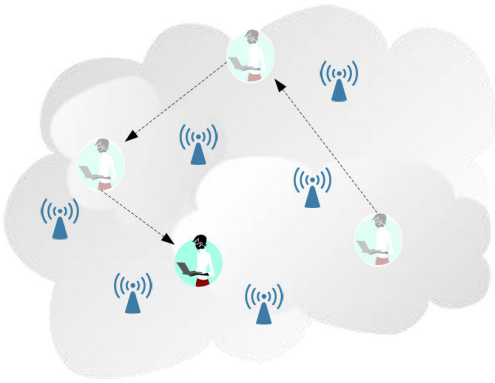


Fig. 2. Moving around to observe surrounding APs.

In the wardriving process, the APs are differentiated by their MAC addresses and indexed from 1 to K, where K is the total number of APs observed in the wardriving process. The observation at location i is recorded in a format as $\mathbf{O}_i = (o_{xi}, o_{yi}, r_{1i}, r_{2i}, \ldots, r_{Ki})$, where (o_{xi}, o_{yi}) is the coordinate of the observing location and $r_{1i}, r_{2i}, \ldots, r_{Ki}$ are the measured signal strength of the APs' beacon signal at (o_{xi}, o_{yi}).

Note that during the listen-only wardriving process, the mobile device does not communicate with the WLAN and just listens to the broadcasted beacon signals. Therefore, the mobile station is transparent to the network and cannot be discovered by an adversary. The information collected in this step is used in step two for estimating the coordinates and the transmit power of the APs. The required effort is part of the cost for protecting location privacy. This cost is acceptable especially when the interested area is a place that the mobile user visits a lot.

B. Estimation of AP locations

Denote (x_k, y_k) as the location coordinate of AP_k and P_{0k} as the signal power level of AP_k at a small reference distance d_0. With the RSS readings of the APs from the wardriving process, the mobile station is able to estimate (x_k, y_k) and P_{0k} for any AP_k in its neighboring area as follows.

Assuming that the beacon signal of an AP_k has the same strength in all horizontal directions, the beacon signal strength of AP_k, the observing locations and the coordinate of AP_k are connected by the following equations according to the log-normal shadowing model [20]:

$$\begin{aligned} r_{k1} &= P_{0k}[(o_{x1} - x_k)^2 + (o_{y1} - y_k)^2]^{-\alpha/2} \\ r_{k2} &= P_{0k}[(o_{x2} - x_k)^2 + (o_{y2} - y_k)^2]^{-\alpha/2} \\ &\cdots \\ r_{kN} &= P_{0k}[(o_{xN} - x_k)^2 + (o_{yN} - y_k)^2]^{-\alpha/2}, \end{aligned} \tag{2}$$

where α is the path loss exponent and N is the number of observing locations. By transforming (2) we can get the equations below.

$$\begin{aligned} (o_{x1} - x_k)^2 + (o_{y1} - y_k)^2 &= \frac{P_{0k}}{r_{k1}}^{2/\alpha} \\ (o_{x2} - x_k)^2 + (o_{y2} - y_k)^2 &= \frac{P_{0k}}{r_{k2}}^{2/\alpha} \\ &\cdots \\ (o_{xN} - x_k)^2 + (o_{yN} - y_k)^2 &= \frac{P_{0k}}{r_{kN}}^{2/\alpha} \end{aligned} \tag{3}$$

Subtracting the last equation from all the other equations, we get an uncorrelated set of equations as follow:

$$\begin{aligned} &2x_k(o_{x1} - o_{xN}) + 2y_k(o_{y1} - o_{yN}) \\ &+P_{0k}^{\frac{2}{\alpha}}(r_{k1}^{-\frac{2}{\alpha}} - r_{kN}^{-\frac{2}{\alpha}}) \\ &\qquad = o_{x1}^2 - o_{xN}^2 + o_{y1}^2 - o_{yN}^2 \\ &2x_k(o_{x2} - o_{xN}) + 2y_k(o_{y2} - o_{yN}) \\ &+P_{0k}^{\frac{2}{\alpha}}(r_{k2}^{-\frac{2}{\alpha}} - r_{kN}^{-\frac{2}{\alpha}}) \\ &\qquad = o_{x2}^2 - o_{xN}^2 + o_{y2}^2 - o_{yN}^2 \\ &\cdots \\ &2x_k(o_{x(N-1)} - o_{xN}) + 2y_k(o_{y(N-1)} - o_{yN}) \\ &+P_{0k}^{\frac{2}{\alpha}}(r_{k(N-1)}^{-\frac{2}{\alpha}} - r_{kN}^{-\frac{2}{\alpha}}) \\ &\qquad = o_{x(N-1)}^2 - o_{xN}^2 + o_{y(N-1)}^2 - o_{yN}^2. \end{aligned} \tag{4}$$

Rewriting (4) into matrix representations, we get

$$\begin{aligned} &\mathbf{A}_k\beta_k = \mathbf{b}_k \\ &\mathbf{A}_k = \begin{bmatrix} 2(o_{x1}-o_{xN}) & 2(o_{y1}-o_{yN}) & r_{k1}^{-\frac{2}{\alpha}} - r_{kN}^{-\frac{2}{\alpha}} \\ 2(o_{x2}-o_{xN}) & 2(o_{y2}-o_{yN}) & r_{k2}^{-\frac{2}{\alpha}} - r_{kN}^{-\frac{2}{\alpha}} \\ \cdots & \cdots & \cdots \\ 2(o_{x(N-1)}-o_{xN}) & 2(o_{y(N-1)}-o_{yN}) & r_{k(N-1)}^{-\frac{2}{\alpha}} - r_{kN}^{-\frac{2}{\alpha}} \end{bmatrix} \\ &\mathbf{b}_k = \begin{bmatrix} o_{x1}^2 - o_{xN}^2 + o_{y1}^2 - o_{yN}^2 \\ o_{x2}^2 - o_{xN}^2 + o_{y2}^2 - o_{yN}^2 \\ \cdots \\ o_{x(N-1)}^2 - o_{xN}^2 + o_{y(N-1)}^2 - o_{yN}^2 \end{bmatrix} \\ &\beta_k = [x_k, y_k, P_{0k}^{\frac{2}{\alpha}}]^T. \end{aligned} \tag{5}$$

Hence, the least square estimation (LSE) of β_k is given by

$$\hat{\beta}_k = [\hat{x}_k, \hat{y}_k, \hat{P}_{0k}^{\frac{2}{\alpha}}]^T = (\mathbf{A}_k^T\mathbf{A})_k^{-1}\mathbf{A}_k^T\mathbf{b}_k. \tag{6}$$

It is important to note that $\hat{P}_{0k}$ is not the exact transmit power of AP_k. It is the estimated received signal strength at a small reference distance d_0 to AP_k.

C. Estimation of path loss to APs

Without loss of generality, suppose the mobile station's desired communication spot is at (o_{x1}, o_{y1}). Based on the estimated AP signal strength level $\hat{P}_{0k}$ from (6), the path loss from AP_k to (o_{x1}, o_{y1}) is calculated by

$$\hat{PL}_k(\text{dBm}) = \hat{P}_{0k}(\text{dBm}) - r_{k1}(\text{dBm}), \quad (7)$$

where r_{k1} is the received beacon signal strength of AP_k at (o_{x1}, o_{y1}).

We use equation (7) to approximate the path loss from AP_k to the mobile user's location (o_{x1}, o_{y1}) and only far-field path loss is considered.

V. Strategy One: Minimizing the Number of RSS Measurements

With the estimated locations of the APs and the path losses between these APs and the mobile station's communication spot (o_{x1}, o_{y1}), we can tune the radiation pattern to limit the number of APs that can hear the mobile station's signal while maintaining the mobile station's communication quality. In the following, we introduce the tuning scheme in two steps. First, the radiation pattern tuning problem is formulated into an optimization problem. Then, this problem is solved to derive the optimal tuning strategy.

A. Tuning problem formulation

To formulate the radiation pattern tuning problem, note that in order to get access to the WLAN, a mobile station has to associate with one of the APs in the WLAN and we denote this AP as AP_c, where $c \in \{1, 2, ...K\}$. For the mobile station to have a stable connection with the WLAN, its signal strength at AP_c must be guaranteed to be larger than the minimum signal strength required by reliable communication, denoted as C_{th}. Hence,

$$P_0|G(\theta_c)|^2(\text{dBm}) \geq \hat{PL}_c(\text{dBm}) + C_{th}(\text{dBm}) + \delta_{\text{dB}}, \quad (8)$$

where δ_{dB} is the maximum error of the path loss estimation, θ_c is the direction of radiation from the mobile station to AP_c, and P_0 is the effective isotropic radiated power at reference distance d_0 from the mobile station. The function $G(\cdot)$ is the beamforming function defined in (1). The radiation angle θ_c towards AP_c can be computed from the estimated coordinate of APs obtained in (6).

While staying connected with AP_c, the mobile station also needs to prevent other APs from detecting its signal. For a particular AP_k, this means that the mobile station needs to guarantee the following situation:

$$\begin{aligned} &\forall \phi_k \in [\theta_k - \delta_\theta, \theta_k + \delta_\theta], \\ &\quad P_0|G(\phi_k)|^2(\text{dBm}) < \hat{PL}_k(\text{dBm}) + R_{th}(\text{dBm}) - \delta_{\text{dB}}, \end{aligned} \quad (9)$$

where R_{th} is the AP receiver sensitivity and δ_θ is the maximum error in the estimation of radiation angle to APs. This error is caused by inaccuracy in AP location estimation.

With the above analysis, the objective of the mobile station, which is preventing too many APs to hear its signal while maintaining stable communications, can be formulated as checking if the following opimtization problem has feasible solutions.

$$\begin{aligned} &\underset{\mathbf{w}}{\text{minimize}}\ \textit{Total Transmit Power} \\ &\text{subject to} \\ &P_0|G(\theta_c)|^2(\text{dBm}) \geq \hat{PL}_c(\text{dBm}) + C_{th}(\text{dBm}) + \delta_{\text{dB}} \\ &y_i = \begin{cases} 1, \text{ if } Z_i \geq R_{th}(\text{dBm}) \\ 0, \text{ otherwise} \end{cases} \\ &Z_i = \max_{\phi_i \in [\theta_i - \delta_\theta, \theta_i + \delta_\theta]} P_0|G(\phi_i)|^2(\text{dBm}) \\ &\qquad\qquad - \hat{PL}_i(\text{dBm}) + \delta_{\text{dB}} \\ &\sum_{i \in \Omega, i \neq c} y_i \leq 2, \end{aligned} \quad (10)$$

where Ω is the set of all the APs that the mobile station can sense in its neighboring area. $\mathbf{w}$ is the complex weight vector of the smart antenna in (1) and can be tuned to change the radiation pattern. Note that we make the constraint that $\sum_{i \in \Omega, i \neq c} y_i \leq 2$ because we need to make sure that the number of APs that can make RSS measurements is smaller than the number required for successful localization. According to the well known trilateration method, if the transmit power of the mobile station is known to the localization system, and the distance between the mobile user and the APs can be estimated trough path loss model, three RSS measurements are required to uniquely localize the mobile station. However in our case, the localization system has no information about the transmit power of the mobile station. This actually adds one unknown variable to the trilateration localization problem. Consequently the required number of RSS measurements for getting a unique location estimation of the mobile station is 4.

B. Solving the tuning problem

The problem formulated in (10) is a mixed-integer nonlinear programming problem and is hard to solve in general. Fortunately, in our situation, it can be easily broken into pure nonlinear programming problems which can be directly solved by existing nonlinear optimization software such like cvx [21] as follows. Since besides AP_c there are $|\Omega| - 1$ APs that may potentially measure the mobile station's signal. To make the number of APs that can take measurements less than 4, we can tolerate at most 2 APs other than AP_c to hear the mobile station's signal, which means at most 2 APs can violate the constraint in (9). Hence, we can convert the feasibility check for (10) into the feasibility check for multiple subproblems, where the constraint in (9) is omitted for 2 APs in each of the subproblem. For example, the subproblem that neglects AP_k and AP_l looks like this:

$$\begin{aligned} &\underset{\mathbf{w}}{\text{minimize}}\ \textit{Total Transmit Power} \\ &\text{subject to} \\ &P_0|G(\theta_c)|^2(\text{dBm}) \geq \hat{PL}_c(\text{dBm}) + C_{th}(\text{dBm}) + \delta_{\text{dB}} \\ &\max_{\phi_i \in [\theta_i - \delta_\theta, \theta_i + \delta_\theta]} P_0|G(\phi_i)|^2(\text{dBm}) \\ &\qquad\qquad - \hat{PL}_i(\text{dBm}) + \delta_{\text{dB}} < R_{th}(\text{dBm}) \\ &\forall i \in \Omega, i \neq c, k, l. \end{aligned} \quad (11)$$

The number of subproblems equals the number of unique combinations of *k* and *l*. If any of the subproblem is feasible, we know that the problem in (10) is feasible and the solution

for that subproblem can be used to tune the radiation pattern. In fact, once we find that for one combination of *k* and *l*, the subproblem in (11) is feasible, we can stop trying other combinations and terminate the optimization process. As a result, in the worst situation, we need to solve $(|\Omega|-1)(|\Omega|-2)/2$ subproblems. In practice $|\Omega|$ is usually not a large number. Hence even solving $(|\Omega|-1)(|\Omega|-2)/2$ subproblems is still acceptable.

Finally if the problem in (10) is infeasible, the number of RSS measurements gathered by the localization system is enough for unique location estimation. In this case, the mobile station switches to the second strategy which is introduced in the next section.

VI. Strategy Two: Maximizing Localization Error

In this strategy, a mobile station aims at maximizing the localization error of a RSS localization system when it cannot be sure that it can limit the number of RSS measurements to a safe level (a.k.a. (10) has no feasible solution). To understand the design of this strategy, in this section, we will first show how we model the problem of maximizing localization bias; then we solve the problem using Genetic Algorithm (GA).

A. Problem Model

Since almost all of the current RSS-based localization schemes are based on the assumption that the mobile station is using omnidirectional antenna, pattern synthesis makes the radiation intensity varies a lot in different directions and therefore introduce large bias in their location estimation.

The problem of maximizing localization bias can be modeled as an optimization problem with the absolute location estimation error as the objective function and the antenna's complex weight vector $\mathbf{w}$ as the variables to be optimized. Meanwhile, the restriction that AP_c is well connected still holds. To formulate the objective function, we first simplify the notation in problem formulation, by setting up a coordinate system, where the location of the mobile station is the origin. Next, we look at how location of the mobile station is estimated by a RSS-based localization system. Denoting R_k as the signal strength of the mobile station at AP_k's location, a RSS-based localization system makes a potentially biased estimation of the mobile station's location, denoted as (x', y'), based on the R_k at all APs. Hence, (x', y') is a function of both AP's location (x_k, y_k) and $R_k, k = 1, 2, ..., K$.

$$(x', y') = F(x_1, \cdots, x_K, y_1, \cdots, y_K, R_1, \cdots, R_K) \quad (12)$$

To maximize the bias in the above estimation, the mobile station computes the location estimation error of the above localization system using the estimated AP coordinates and the predicted RSS at the APs. Based on the estimation of its path loss to the APs in Section IV-C, the mobile station can estimate its signal strength at AP_k as:

$$\hat{R}_k(\text{dBm}) = P_0|G(\theta_k)|^2(\text{dBm}) - \hat{PL}_k(\text{dBm}). \quad (13)$$

In addition, from Section IV-B, mobile station can also estimate the coordinate of AP_k as at $(\hat{x}_k, \hat{y}_k)$. Based on these information, the mobile station can make an estimation of (x', y'), which is the RSS-localization system's estimate of its location. Denote $(\hat{x}', \hat{y}')$ as the mobile station's estimate of (x', y'). For the mobile station, the optimization problem of maximizing the bias of a RSS-localization system can be modeled as

$$\begin{aligned}
&\underset{\mathbf{w}}{\text{maximize}} \;\; E[||(\hat{x}', \hat{y}')||] \\
&\text{subject to} \\
&\quad P_0|G(\theta_c)|^2(\text{dBm}) \geq \hat{PL}_c(\text{dBm}) + C_{th}(\text{dBm}) + \delta_{\text{dB}} \\
&\quad (\hat{x}', \hat{y}') = F(\hat{x}_i, \hat{y}_i, \hat{R}_i) \\
&\quad \hat{R}_i(\text{dBm}) = P_0|G(\gamma_i)|^2(\text{dBm}) - \hat{PL}_i(\text{dBm}) \\
&\quad \gamma_i \sim U[\theta_i - \delta_\theta, \theta_i + \delta_\theta] \\
&\quad (\text{i.e. } \gamma_i \text{ is uniformly distributed in } [\theta_i - \delta_\theta, \theta_i + \delta_\theta]) \\
&\quad i = 1, 2, ..., K.
\end{aligned} \quad (14)$$

To solve the above optimization problem, the location estimation function $F(\cdot)$ must be known. In the remainder of this paper, we assume that the localization system uses least square (LS) method since it is a very popular scheme used in many localization systems [22]. Under this method, the function $F(\cdot)$ can be deduced from (5).

It is important to note that our scheme is not limited to least square method. If the mobile station knows that another localization algorithm is used by the RSS-based localization system, the mobile station can change function $F(\cdot)$ and optimize its radiation pattern accordingly. Many may argue that the estimation algorithm adopted by the localization system may not be the common least square method (e.g. a fingerprint based estimation algorithm) and the algorithm is not known to the mobile station. In this case, the mobile station has no certain ways to maximize the bias of the location system. Nevertheless, we believe that despite the wrong guess on location estimation algorithm, the irregular radiation pattern produced by the optimization problem in (14) should at least increase the bias for any estimation methods even it cannot maximize it.

B. Genetic Algorithm (GA) Solution

GA is used to solve the problem in (14) since it does not require derivative information of the objective function and it works without limit on the number of variables [23], [24]. The GA procedure in this paper is similar to its standard form and small modification is made to fit it to our problem.

1) Chromosome Construction: In our application of pattern synthesis, the complex weight vector $\mathbf{w} = [w_1, w_2, \cdots, w_{N_{ant}}]^T$ is directly used as a chromosome. Each element in $\mathbf{w}$ represents an excitation factor of the corresponding antenna element. Therefore the length of the chromosome is determined by the number of antenna elements N_{ant}.

2) Initial Population: The number of the chromosomes in the population is denoted by N_{pop}. So the population is actually an $N_{pop} \times N_{ant}$ complex matrix. We initialize the population by randomly generate the real and complex part of the complex gene as a random number within $[0, \frac{1}{N_{ant}}]$.

3) Natural Selection: Each chromosome in the population is passed to the objective function to calculate corresponding output objective value. Then the chromosomes are sorted

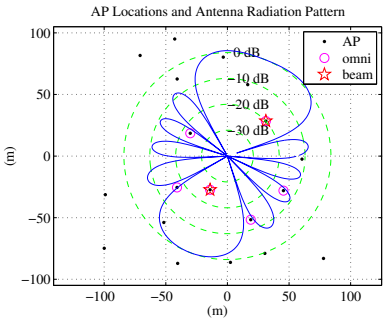


Fig. 3. An beamforming pattern that limits the number of APs in range to be less than 4. (The radiation pattern of the omnidirectional antenna is the out most green circle.)

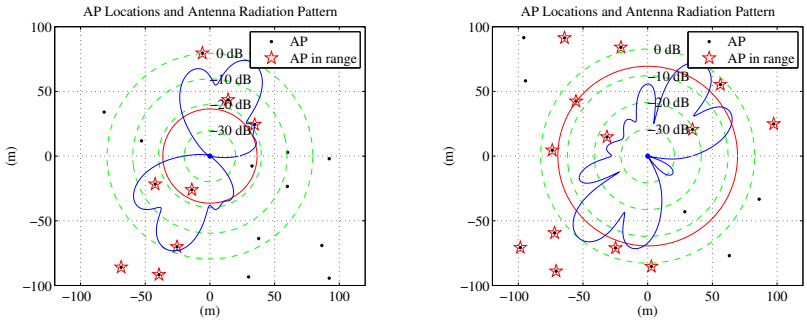


Fig. 4. Two examples of the optimized beamforming pattern.

according to descending order of their associated objective values. We just keep the best N_{ns} chromosomes and discard the others.

4) Mating: Mate selection is a procedure that two parent chromosomes are picked to produce offspring chromosomes. In our problem, the chromosomes with greater objective value should have higher chance to be selected. We use roulettle wheel selection method [23] and assign higher probability to high ranking chromosomes.

5) Reproduction: Decimal crossover is used to produce offspring chromosomes. Given the parent chromosomes $\mathbf{w}_f$ and $\mathbf{w}_m$, the child chromosomes are given by:

$$\begin{aligned} \text{child1} &= 0.5\mathbf{w}_f + 0.5\mathbf{w}_m \\ \text{child2} &= 1.5\mathbf{w}_f - 0.5\mathbf{w}_m \\ \text{child3} &= -1.5\mathbf{w}_f + 0.5\mathbf{w}_m. \end{aligned} \tag{15}$$

This makes the offspring unbounded by the parent chromosomes while keeps good property of the parent chromosomes.

6) Mutation: In order to keep the best chromosome in the population, the top ranking chromosome will not be mutated. According to the mutation rate, a group of the genes in the rest chromosomes are randomly selected and their values are reset as a random number within $[0, \frac{1}{N_{ant}}]$. This process is helpful to prevent GA evolution from stagnating at a local optimum.

7) Terminating Criteria: The evolution process is repeated until a maximum total number of generations is reached. When GA is terminated, the complex weight vector associated with the optimal objective value is adopted in radiation pattern synthesis. The output optimal objective value is the predicted location estimation bias of the localization system.

Note that due to the possible error in estimating the APs' locations and their angles to the mobile station, difference between the predicted localization estimation bias and the true location estimation bias might exist. However, decision making based on the existing information is the best thing that the mobile user can do.

VII. Simulation Results

To evaluate the performance of our proposed location privacy scheme, we simulate a WLAN and our privacy protection scheme in Matlab environment. In our simulation, APs are randomly deployed in a 200×200 m^2 2-D space. They are spaced reasonably far away from each other to mimic real network deployment since real WLAN deployment rarely have two APs sit very close to each other to avoid interference and increase efficiency in network coverage.

We first show examples of individual synthesized beamforming patterns of both the first and the second strategy. Then we use Monte Carlo simulation to estimate the success rate of the first strategy and the expected localization bias caused by the second strategy with different combinations of the number of APs and the number of antenna elements.

The setting of the parameters is as follows. The path loss exponent α is set to be 3, $C_{th}(\text{dBm}) = -75$ dBm and $R_{th}(\text{dBm}) = -80$ dBm. Transmit power of the APs are randomly generated in the range of 10 dBm to 20 dBm. For both omnidirectional antenna cases and pattern synthesis cases, shadowing noise variance is 1. For the proposed scheme, $\delta_{\text{dB}} = 5$ dBm and $\delta_\theta = 5^\circ$. Following is the detailed discussion of the simulation results.

A. Examples of Synthesized Patterns

In the following examples, the linear antenna array consists of 6 isotropic elements and the interval spacing between neighboring antenna elements is half the wavelength.

1) Strategy one: limiting AP measurements to be less than 4: Fig. 3 illustrates an example of synthesized radiation pattern generated by strategy one of our privacy protection scheme. In this figure, the mobile station is at the origin. When the privacy protection scheme [13] based on power control over omnidirectional antenna is used to protect user privacy, the APs that can hear the mobile station's signal are marked by magenta circles and there are 6 of them. While using the optimized radiation pattern under our strategy one, the APs that can hear the mobile station's signal are marked by red pentagrams and there are only 2 such APs. For both situations, the transmit power is chosen to be the minimum value that satisfies the communication requirement defined by (8).

From the above example, it is easy to see that if the mobile station uses omnidirectional antenna, it will be heard by 6 APs and it can be localized. While by using our strategy one of radiation pattern synthesis, the mobile station is able to

concentrate the transmit power for communication purpose and limit the power emitted to other directions. Therefore the number of APs that can hear its signal is significantly reduced.

2) Strategy two: maximizing the location estimation bias: When it is inevitable that more than 4 APs can get RSS measurements of its signal, the mobile station optimizes its radiation pattern with GA to bias the location estimation as much as it can. In the GA compuation process, the maximum number of generation is set as 500.

Fig. 4 illustrates two examples of the optimized radiation pattern overlapping with the AP distribution in the 2-D space. The radius of red circle in the center is the size of estimation error of LS location estimation algorithm. As we can see from Fig. 4(a), when the localization error is maximized, our strategy controls the APs that can measure RSS to be on a diagonal that passes the location of the mobile station. This observation is in accordance with the result in [25], which says that collinear deployment of anchors has negative impact on localization performance. In Fig. 4(b), the directional gain of the smart antenna varies a lot in different direction. Consequently even if the localization system successfully get abundant RSS measurements of the mobile station, those RSS measurements are biased. Directly using these RSS measurements to do localization will cause large estimation error, while refining those measurements are very difficult without knowing the radiation pattern used by the mobile station.

Essentially, we can see that our strategy two degrades the localization performance by not only distorting the RSS measurement but also selectively only letting the APs located collinearly measure RSS.

B. Statistical analysis

In this part, we first use Monte Carlo simulation to estimate the success rate of our strategy one, which attempts to limit the number of APs that can hear the mobile station's signal to be below 4. Then, we use Monte Carlo simulation to evaluate the localization bias introduced by our strategy two. We evaluate our scheme under different combinations of the number of APs (N_{ap}) and the number of antenna elements (N_{ant}). For each pair of N_{ap} and N_{ant}, 200 simulation runs are conducted.

1) Success rate of the first-priority strategy: Table I lists the success rate of the our strategy one. $N_{ant} = 1$ is actually the omnidirectional antenna case and the data in the first row shows the simulated success rate of using transmit power control to limit the number of APs in range to be less than 4. The higher success rate of our pattern synthesis scheme compared with the omnidirectional antenna case indicates that pattern synthesis is more powerful to reduce the number of RSS measurements while keeping a mobile station's communication quality intact.

Generally, the success rate tends to increase when N_{ant} increases since smart antenna array with more antenna elements has more flexibility in tuning the radiation pattern. The capability of concentrating the transmit power to desired direction reduces the unnecessary power emission to other directions and keep the mobile station more secure.

TABLE I
SUCCESS RATE OF STRATEGY ONE

N_{ant}	N_{ap} = 10	N_{ap} = 20	N_{ap} = 30
1	0.335	0.355	0.405
4	0.51	0.455	0.595
6	0.575	0.59	0.615
8	0.615	0.61	0.615
10	0.635	0.64	0.7
12	0.705	0.655	0.715

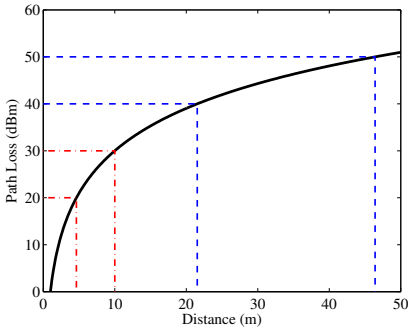
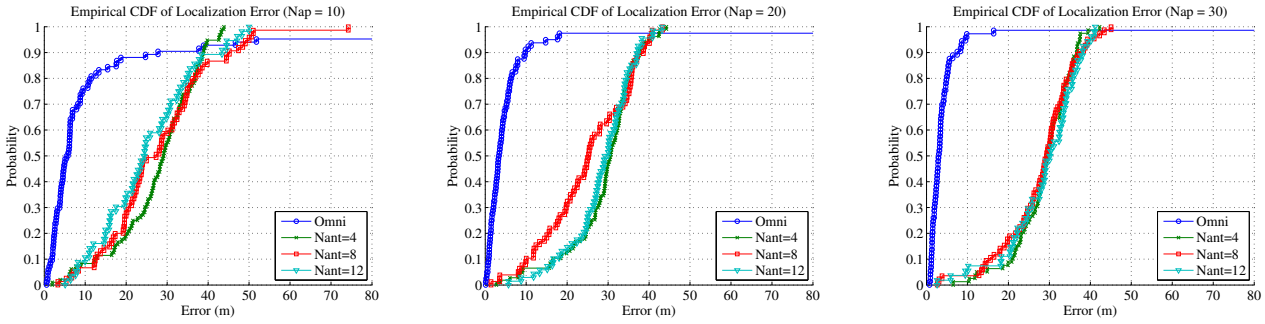


Fig. 5. Large scale path loss v.s. distance using log-normal shadowing model with $\alpha = 3$.

We also find that when N_{ap} increases, the success rate of our strategy one also may increase. This is counterintuitive in a sense that higher density of APs should lead to greater chance that the signal from the mobile station to be heard by more APs. However, note that, due to variations in path loss and inherent noise in our AP location estimation, APs in areas that are a little further away than AP_c may still hear the mobile station's signal. We call this as the blurring zone around AP_c and our strategy one has conservatively considered APs in the bluring zone as APs that can measure RSS of the mobile station. According to log-normal shadowing path loss model illustrated by Fig. 5, the further away that AP_c is from the mobile station, the more flat the curve of signal attennuation is, which results in a larger blurring zone around AP_c. When AP density is high, AP_c usually locates closer to the mobile station, resulting a small bluring zone around AP_c. As the blurring zone is smaller, there is less chance that any AP will happen to be in this blurring zone. Hence, the success rate of strategy one increases.

2) Localization bias introduced by strategy two: Fig. 6 shows the cumulative probability distribution (CDF) of the localization bias caused by our strategy two. As we can see from the figure, when omnidirectional radiation pattern is used, the localization bias is around 10 meters and the bias is mainly caused by noise. Comparatively, the CDF curves of our strategy two are more on the right side, indicating significantly larger location estimation bias. Consider that the radius of an AP's coverage area is usually 30 to 50 meters, this estimation bias is fairly large. Thus we can draw the conclusion that our pattern synthesis-based privacy protection scheme can bring large bias to the location estimation result



(a) (b) (c)

Fig. 6. CDF of localization error caused by pattern synthesis.

of localization system. However, we do not observe big performance difference among antenna arrays with different number of antenna elements. Hence we conclude that antenna array with different number of antenna elements have similar performance in distorting the localization result.

VIII. Conclusion

In WLAN environment, a mobile station's location can be revealed by its radiated signal and this poses great threat to the mobile user's location privacy. In this paper, we proposed a physical layer location privacy protection scheme against RSS-based localization systems using a linear smart antenna array on the mobile station. Our scheme consists of a passive observation step and an antenna pattern synthesis step. In the passive observation step, the mobile station moves around to measure the beacon signal of the neighboring APs and estimate the path losses to the APs. In the pattern synthesis step, two pattern optimization strategies are used to prevent the positioning system from correctly localizing the mobile station while the communication quality is not affected. Simulation results show that our first-priority strategy significantly increases the probability of inadequate RSS measurements and the second strategy substantially degenerates the localization precision.

Acknowledgment

This work is supported by National Science Foundation fund number ECCS-0802112 and the Institute for Critical Technology and Applied Science (ICTAS).

References

[1] B. Schilit, J. Hong, and M. Gruteser, "Wireless location privacy protection," *Computer*, vol. 36, pp. 135–137, 2003.

[2] C. Tang and D. O. Wu, "Mobile privacy in wireless networks-revisited," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 1035–1042, Mar. 2008.

[3] M. Robinson and I. Psaromiligkos, "Received signal strength based location estimation of a wireless lan client," in *Proceedings of IEEE WCNC*, vol. 4, 2005, pp. 2350–2354.

[4] C. Feng, W. Au, S. Valaee, and Z. Tan, "Compressive sensing based positioning using rss of wlan access points," in *Proceedings of IEEE INFOCOM*, 2010.

[5] P. Bahl and V. N. Padmanabhan, "Radar: an in-building rf-based user location and tracking system," in *Proceedings of IEEE INFOCOM*, vol. 2, 2000, pp. 775–784.

[6] S. Kim, H. Jeon, and J. Ma, "Robust localization with unknown transmission power for cognitive radio," in *Proceedings of IEEE MILCOM*, 2007, pp. 1–6.

[7] L. Huang, K. Matsuura, H. Yamane, and K. Sezaki, "Enhancing wireless location privacy using silent period," in *Proceedings of IEEE WCNC*, vol. 2, 2005, pp. 1187–1192.

[8] P. Bahl and V. N. Padmanabhan, "Enhancements to the radar user location and tracking system," Microsoft Research, Tech. Rep., 2000.

[9] K. Kaemarungsi and P. Krishnamurthy, "Modeling of indoor positioning systems based on location fingerprinting," in *Proceedings of IEEE INFOCOM*, vol. 2, 2004, pp. 1012–1022.

[10] A. M. Ladd, K. E. Bekris, A. Rudys, G. Marceau, L. E. Kavraki, and D. S. Wallach, "Robotics-based location sensing using wireless ethernet," *Wireless Networks*, vol. 11, pp. 189–204, 2005.

[11] M. Decker, "Location privacy-an overview," in *Proceedings of the 2008 7th International Conference on Mobile Business ICMB '08*, 2008, pp. 221–230.

[12] J. Meyerowitz and R. Roy Choudhury, "Hiding stars with fireworks: location privacy through camouflage," in *Proceedings of ACM MOBICOM*, 2009, pp. 345–356.

[13] T. Jiang, H. J. Wang, and Y.-C. Hu, "Preserving location privacy in wireless lans," in *Proceedings of ACM MOBISYS*, 2007, pp. 246–257.

[14] F. L. Wong, M. Lin, S. Nagaraja, I. Wassell, and F. Stajano, "Evaluation framework of location privacy of wireless mobile systems with arbitrary beam pattern," in *Fifth Annual Conference on Communication Networks and Services Research (CNSR '07)*, 2007.

[15] Y. Chen, K. Kleisouris, X. Li, W. Trappe, and R. P. Martin, "The robustness of localization algorithms to signal strength attacks: a comparative study," in *Proc. Intl Conf. Distributed Computing in Sensor Systems (DCOSS)*, 2006.

[16] K. Bauer, D. McCoy, E. Anderson, M. Breitenbach, G. Grudic, D. Grunwald, and D. Sicker, "The directional attack on wireless localization: how to spoof your location with a tin can," in *Proceedings of IEEE GLOBECOM*, Nov. 2009, pp. 1–6.

[17] H. Lebret and S. Boyd, "Antenna array pattern synthesis via convex optimization," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 526–532, Mar. 1997.

[18] D. Han, D. G. Andersen, M. Kaminsky, K. Papagiannaki, and S. Seshan, "Access point localization using local signal strength gradient," in *Proceedings of Passive & Active Measurement (PAM)*, Seoul, South Korea, Apr. 2009.

[19] kismet, *http://www.kismetwireless.net/index.shtml*.

[20] T. S. Rappaport, *Wireless Communications: Principles & Practice*. Prentice Hall, 2002.

[21] M. Grant and S. Boyd, "Cvx: Matlab software for disciplined convex programming, version 1.21." http://cvxr.com/cvx, May 2010.

[22] J. Yang and Y. Chen, "Indoor localization using improved rss-based lateration methods," in *Proceedings of IEEE GLOBECOM*, 2009, pp. 1–9.

[23] R. L. Haupt and D. H. Werner, *Genetic Algorithms in Electromagnetics*. A John Wiley & Sons, Inc., 2007.

[24] K.-K. Yan and Y. Lu, "Sidelobe reduction in array-pattern synthesis using genetic algorithm," *IEEE Transactions on Antennas and Propagation*, vol. 45, pp. 1117–1122, 1997.

[25] Y. Chen, J.-A. Francisco, W. Trappe, and R. Martin, "A practical approach to landmark deployment for indoor localization," in *Proceedings of IEEE SECON*, vol. 1, 2006, pp. 365–373.